

REMARKS

In the Official Action mailed on **23 August 2006**, the Examiner reviewed claims 1-62. Claims 1-62 were rejected under 35 U.S.C. §101 because the claims are directed to non-statutory subject matter. Claims 1-62 were rejected under 35 U.S.C. §112, second paragraph as being indefinite. Claims 1-3, 6, 21-23, 26, 41-43, 46, and 61-62 were rejected under 35 U.S.C. §102 as being anticipated by Dillion et al (USPub 2003/0088562, hereinafter “Dillion”).

Rejections under 35 U.S.C. §101

Claims 1-62 were rejected because the claims are directed to non-statutory subject matter.

Applicant has amended independent claims 1, 21, and 41 to include two new limitations: (1) “wherein the set of components provides an abstract representation for the document;” and (2) “wherein the abstract representation is subsequently used as a substitute for the document during query operations involving the document.” These added limitations address each of the following questions.

(1) What is the practical application?

For a given textual document such as a web page, the instant application provides an abstract representation for the given document, which is a vector containing a set of components. In one embodiment of the present invention, each component contains a probabilistic numerical value. This representation is compact while capturing the semantic meaning or topic associated with the text in the document.

(2) What is the final result that is concrete, useful and tangible?

The abstract representation of the document can be used in many applications related to the document. For example, to respond to a query, the system compares the abstract representation of the document with the query,

instead of comparing the original textual document with the query. This can be more accurate and more efficient in generating query result.

Consequently, the new limitations provide patentable subject matter. These amendments find support in paragraphs [00217] – [00218], and [00226] – [00234] (for the first limitation), and paragraph [00201] (for the second limitation) of the instant application.

Claims 61-62 are rejected because they are not limited to tangible embodiments. Applicant has amended paragraph [0055] to remove the transmission media and related embodiments. As a result of this amendment, claims 61-62 are now only limited to statutory subject matter of storage media.

Further in regarding to claims 61-62, Applicant respectfully points out that the claims are directed to claiming a computer-readable storage medium containing a data structure, but not the data structure itself. A computer-readable storage medium is a tangible and patentable subject matter.

Rejections under 35 U.S.C. §112

Claims 1-62 were rejected under 35 U.S.C. §112, second paragraph as being indefinite. Applicant has modified the limitation from “wherein the conceptually related words are words that relate to a single idea,” to “wherein the conceptually related words are words that relate to a common topic.” As an example of a cluster of conceptually related words, [*career jobs employment recruitment part-time salary*] are conceptually related words to the topic of “job”. This amendment finds support in [00187] – [00190] of the instant application.

Rejections under 35 U.S.C. §102

Independent claims 1, 21, 41, and 61 were rejected as being anticipated by Dillion. Applicant respectfully points out that Dillion teaches analyzing a textual document by generating a **list of weighted keywords**. More specifically, the

textual document is first parsed into words, wherein keywords are extracted, and the weight of each unique keyword is computing (see Dillion, paragraph [0065]).

In contrast, the present invention characterizes a textual document by constructing a set of components for the document, wherein each component indicates the **likelihood (or the probability) of the document is related to a candidate cluster of conceptually related words**. Note that the cluster conceptually related words are **words that relate to a common topic**. For example, cluster [*career jobs employment recruitment part-time salary*] contains conceptually related words to the topic of job (see paragraphs [00187]-[00190] of the instant application). Note that there exist millions of topics. However, the set of components for the document only contain a small subset of these possible topics (see FIG. 20), each are referred to as candidate clusters. These candidate clusters have higher probability of generating the document (see [00216]).

Characterizing a document this way is beneficial because (1) it can effectively capture the underlying semantic meaning in the text of the document; (2) it provides an abstract representation for the document, which facilitates performing searching operations related to the document more efficiently and accurately. There is nothing within Dillion, either explicit or implicit, which suggests relating a document to a set of conceptually related words, wherein the conceptually related words relate to a common topic.

Accordingly, Applicant has amended independent claims 1, 21, 41, and 61 to clarify that the present invention characterizes a textual document by constructing a set of components, wherein each component indicates the likelihood (or the probability) of the document is related to a candidate cluster of conceptually related words. These amendments find support in paragraphs [00187]-[00190], [00215]-[00218], and [00226] – [00231] of the instant application.


Hence, Applicant respectfully submits that independent claims 1, 21, 41, and 61 as presently amended are in condition for allowance. Applicant also

submits that claims 2-20, which depend upon claim 1, claims 22-40, which depend upon claim 21, claims 42-60, which depend upon claim 41, and claim 62, which depends upon claim 61, are for the same reasons in condition for allowance and for reasons of the unique combinations recited in such claims.

CONCLUSION

It is submitted that the present application is presently in form for allowance. Such action is respectfully requested.

Respectfully submitted,

By 
Edward J. Grundler
Registration No. 47,615

Date: 23 October 2006

Edward J. Grundler
PARK, VAUGHAN & FLEMING LLP
2820 Fifth Street
Davis, CA 95618-7759
Tel: (530) 759-1663
FAX: (530) 759-1665
Email: edward@parklegal.com